# Analyzing RGB Images using Topology: How to use discrete Morse theory to study crime data

Chuan Du, Adarsh Manawa, Christopher Szul, Nima Rasekh, Ruth Davidson, Rosemary Guzman

University of Illinois at Urbana-Champaign

January 10, 2017

# How our project came to be: IGL

This project started as part of the *Illinois Geometry Lab* (IGL) at the University of Illinois. It teams up several undergraduate students to work under the supervision of faculty.

- *Faculty*: Ruth Davidson, Rosemary Guzman
- *Graduate Student*: Nima Rasekh
- *Undergraduate Students*: Chuan Du, Adarsh Manawa, Christopher Szul

# Background: Work at ANU

Our work is based on:

- **Paper**: "Skeletonization and Partitioning of Digital Images Using Discrete Morse Theory"
- **Authors**: *Olaf Delgado-Friedrichs*, *Vanessa Robins*, and *Adrian Sheppard* from Australian National University (ANU)
- **Published**: IEEE transactions on pattern analysis and machine intelligence, vol. 37, no. 3, pp. 654-666, 2015.
- **Code**: https://github.com/AppliedMathematicsANU/diamorse

In that paper the authors use discrete Morse theory to analyze grayscale images.

# Overview of ANU work

1. Input a given grayscale image.
2. Use adjacency of pixels to form a cubical complex, where each point represents a pixel.
3. Use grayscale values on points to build discrete Morse function.
4. Use discrete Morse function to compute persistent homology.
5. Use persistent homology to gain information about image.

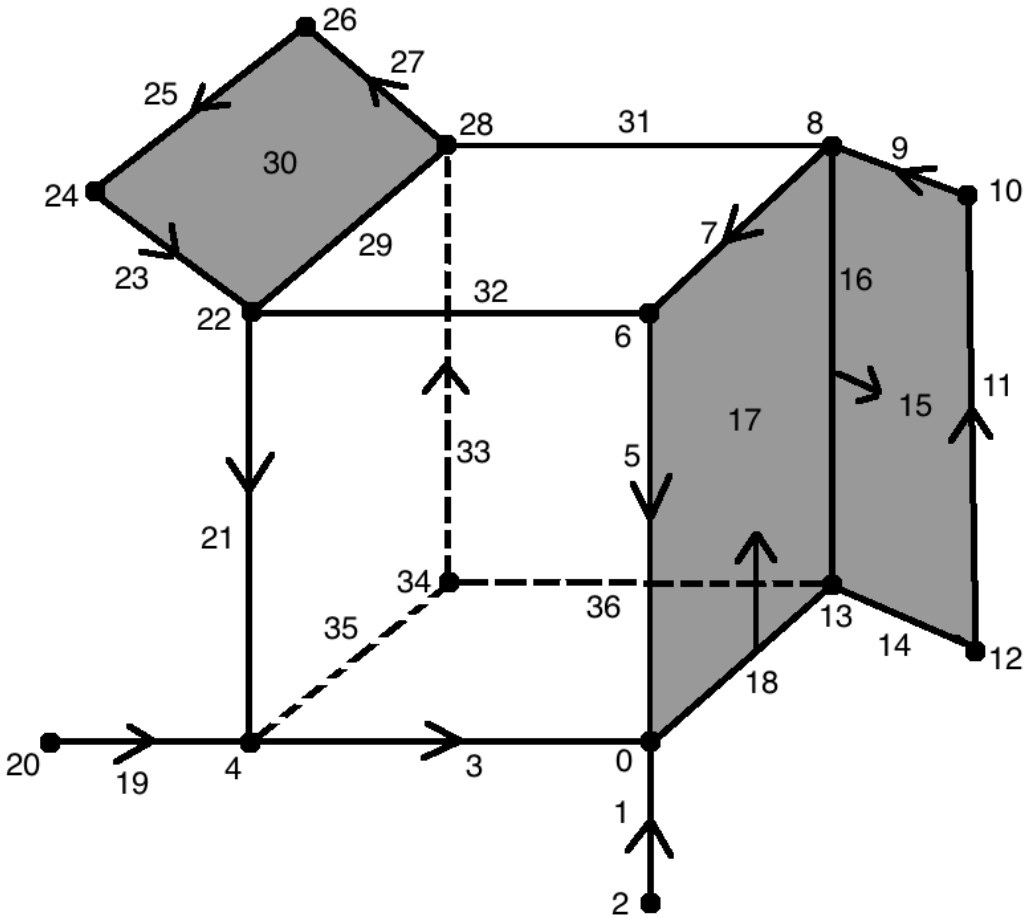
# Grayscale values

A grayscale image is a collection of pixels where each pixel has a gray scale value attached to it:

- **0**: Black
- **1** - **254**: Shades of Gray
- **255**: White

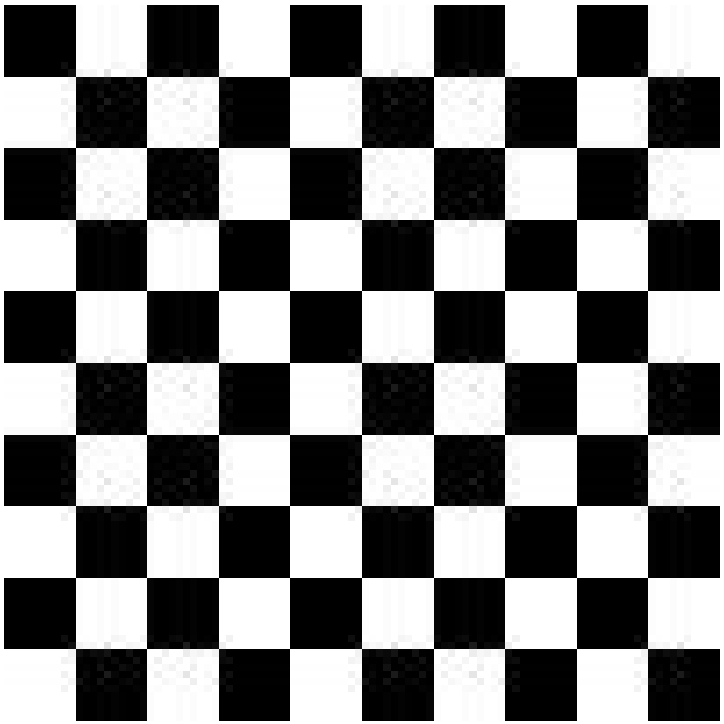
# Cubical Complex

The code transforms such grayscale image into a cubical complex with a discrete Morse function

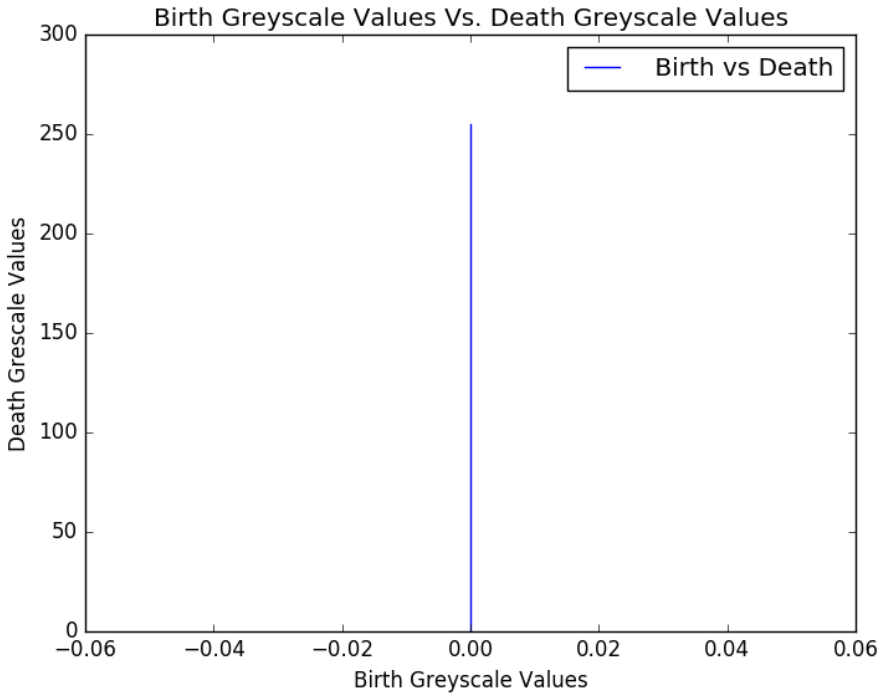


# First Example

## Sample Image



## Persistent Barcode



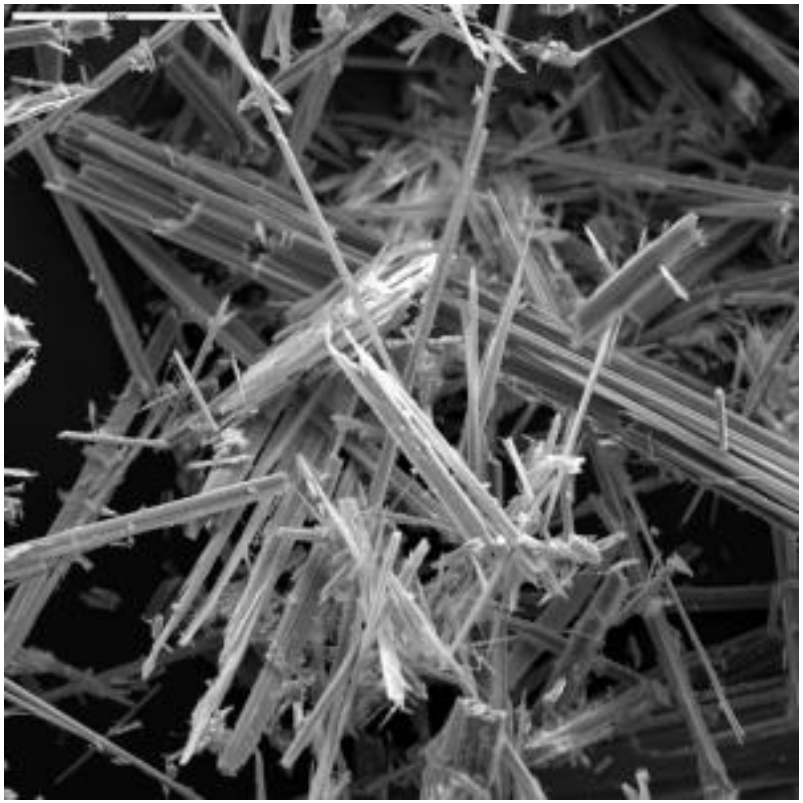
# What does any of this mean?

- **Existence**: The existence of non-trivial cycles indicates that there is a change in color happening in the image.
- **Persistence**: The persistence of each cycle gives us the contrast of the change. The higher the contrast in color, the more persistent the cycle.

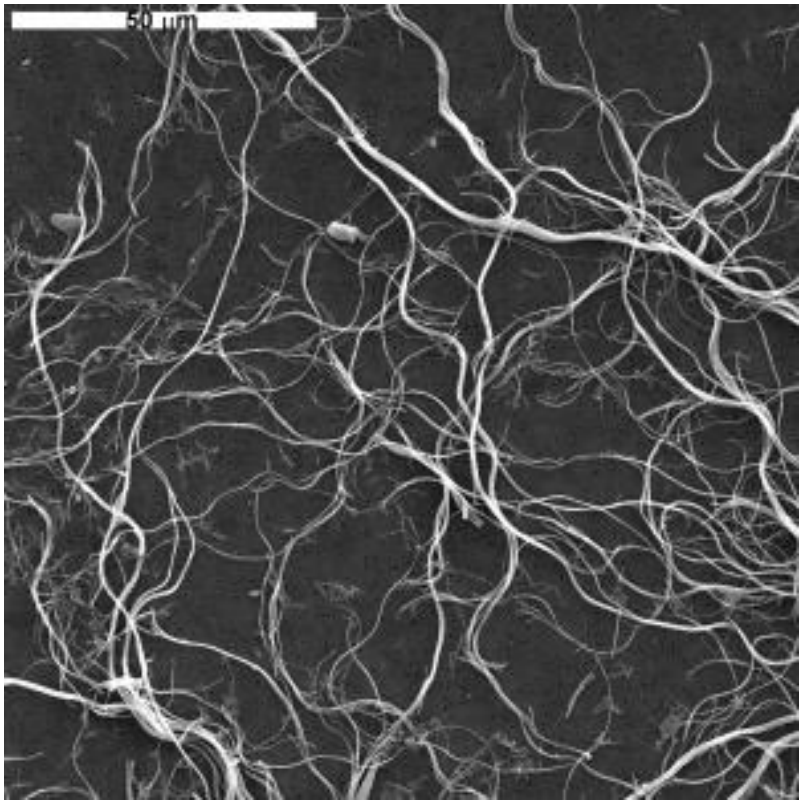
# Practical Applications: Analyzing Asbestos

Our first work as a team was to compute some examples:

Anthophyllite Asbestos

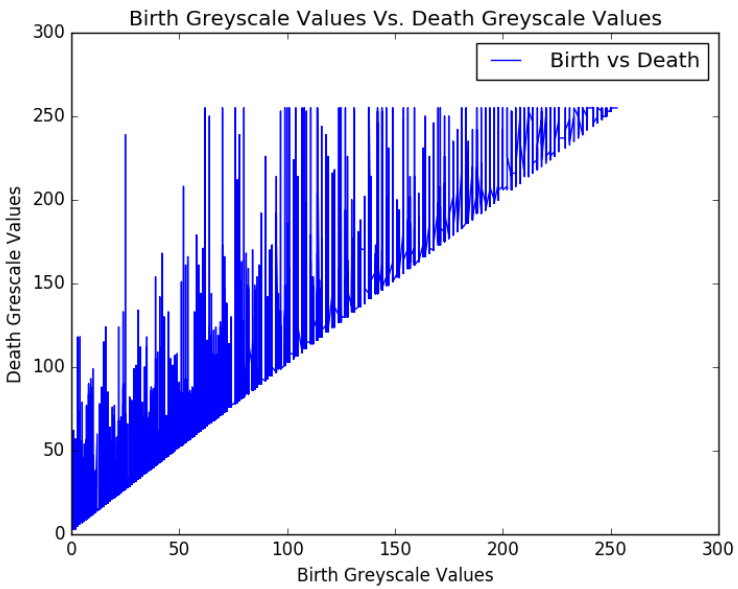


Chrysotile Asbestos

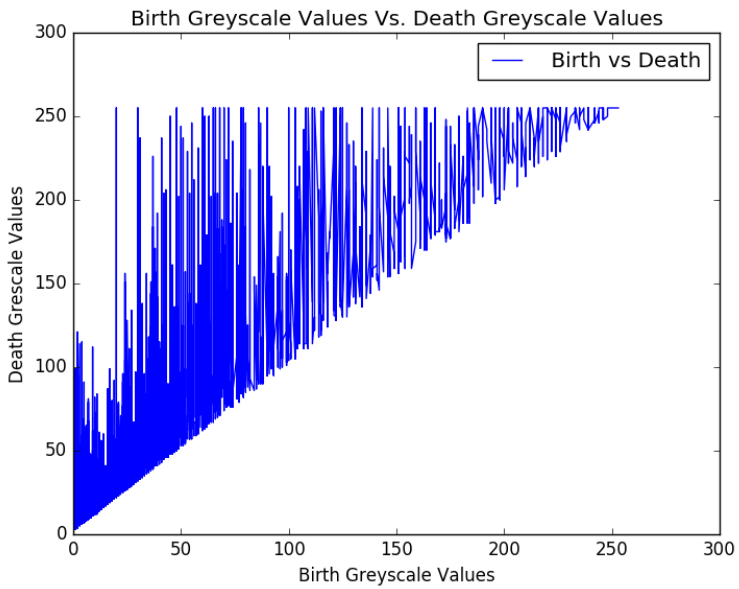


# Practical Applications: Analyzing Asbestos

## Anthophyllite Barcode




## Chrysotile Barcode



# Expand from Grayscale to RGB

An RGB image assigns to each pixel a 3-tuple determining the shade of red, green and blue. Our goal was to expand the functionality of the code so that it can also analyze RGB images and then use that code on various heat maps that can help us do data analysis.
All work described from here on can be found in the paper:

- **Title**: "RGB image-based data analysis via discrete Morse theory and persistent homology "
- **Authors**: Chuan Du, Christopher Szul, Adarsh Manawa, Nima Rasekh, Rosemary Guzman, and Ruth Davidson
- **Link**: https://faculty.math.illinois.edu/∼rasekh2/dmt_vf_ms.pdf
- **Code**: https://github.com/redavids/IBTCDA/tree/master

# How to get from RGB to Grayscale

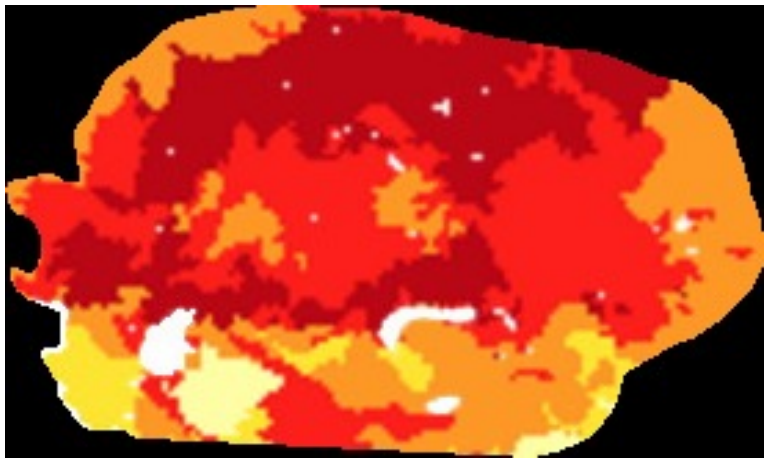
We used 3 methods to convert RGB to grayscale.

1. *Average Method*: Take average of three RGB values.
2. *Luminosity Method*: Use the formula $0.21R + 0.72G + 0.07B$.
3. *Custom Convertio*: Use the online converter "Convertio" (see https://convertio.co/).

We use different methods because colors can have statiscal meaning which can be lost using only one converter.

# First Application: Variability of Water Scarcity

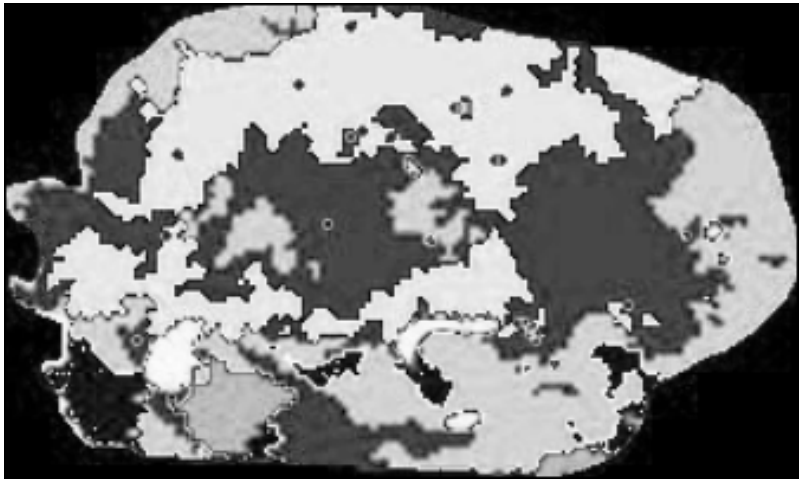
This is a heat map of of *Kazakhstan* indicating regional water scarcity levels. This open-source map can be found at http://bit.ly/2fEoU7Q.



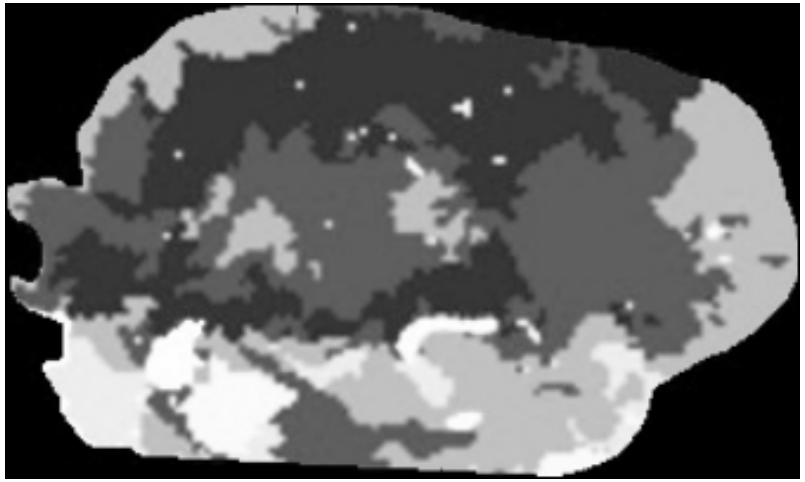
We use the three methods explained above to show how we can extract topological data.

# 3 Ways of Converting to Grayscale

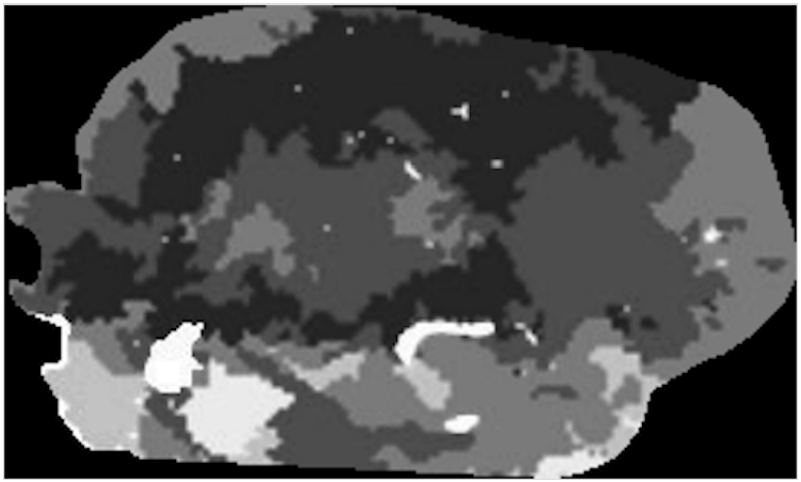
Average



Luminosity

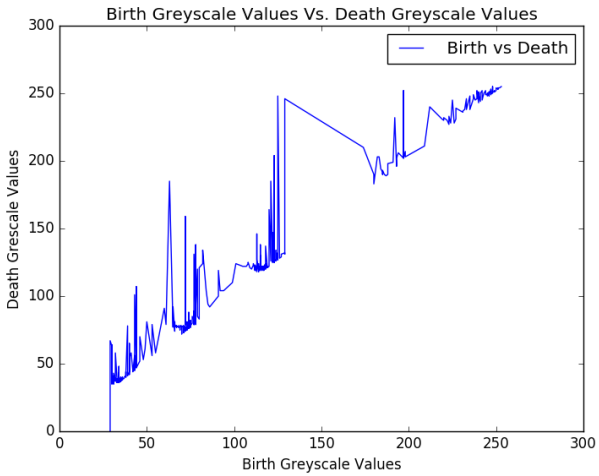


Convertio

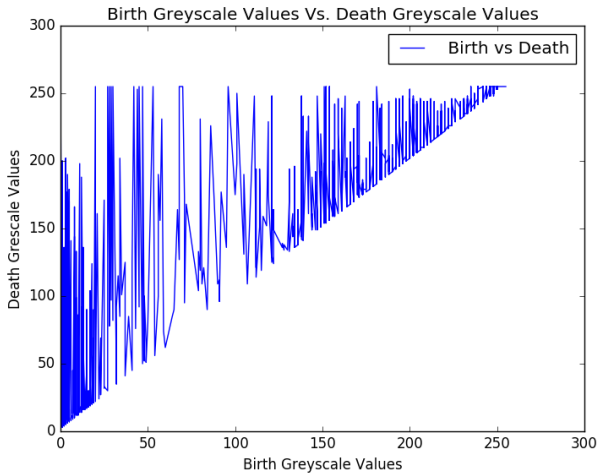


# 3 Ways to get Barcodes

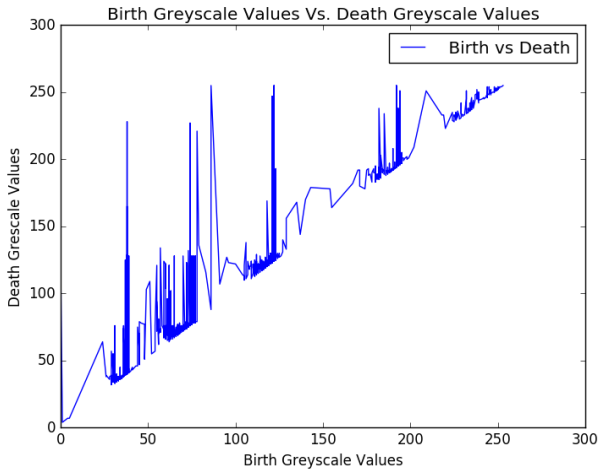
## Average




## Luminosity




## Convertio



# Different methods, different results

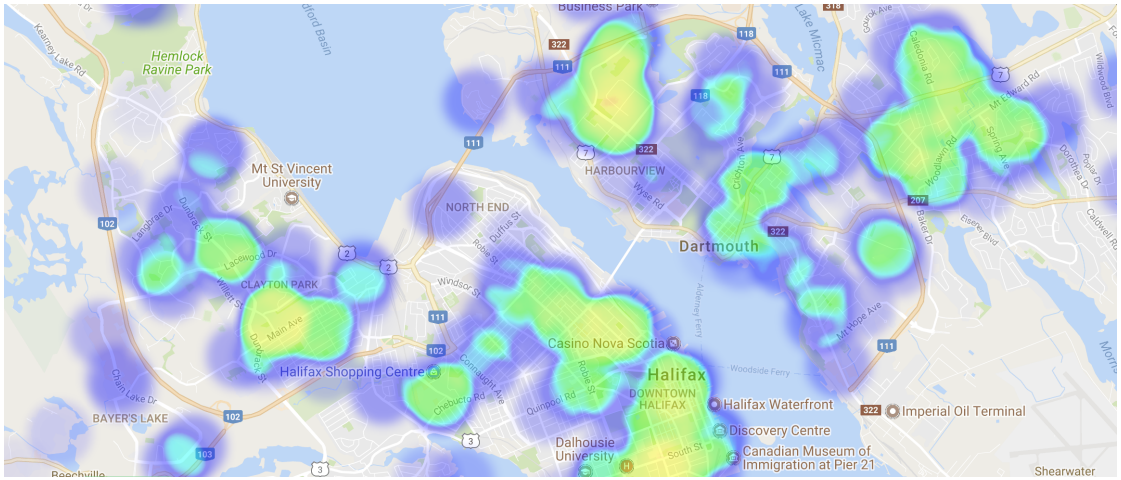
As we see from these results different methods give us different barcode diagram. Depending on the kind of information we want to focus on we have to be able to calibrate the RGB-to-grayscale conversion to our needs.

# Second Application: Crime Data Analysis

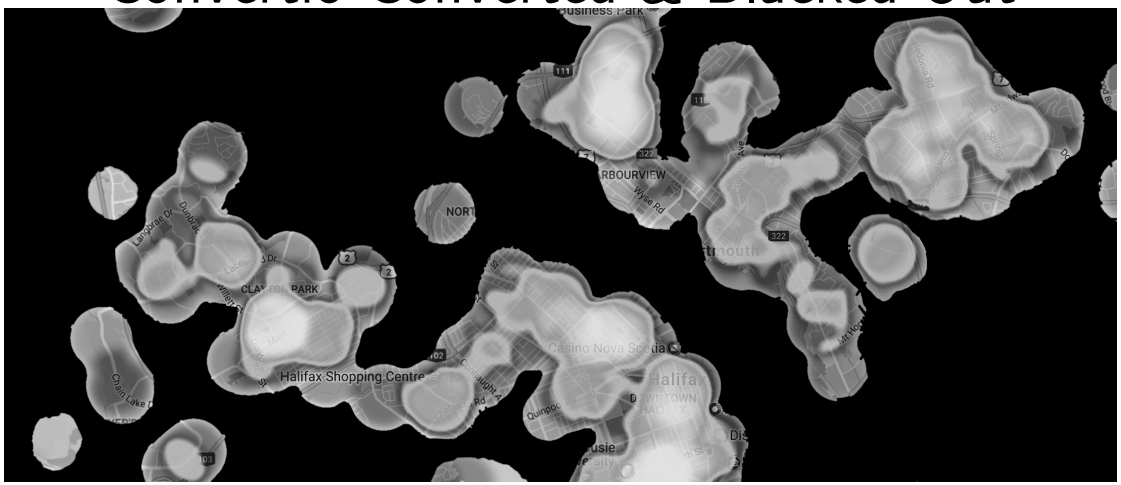
- **Goal**: Predictive analysis of crime pattern via comparion of persistent barcodes.
- **Source**: Halifax, Nova Scotia Crime maps powered by OpenDataHalifax (see http://www.crimeheatmap.ca/).
- **Dates**: Feb 1st, 2017 to April 12th, 2017 (6 maps).

# Crime Heat Map for Feb 1st

## Original Image

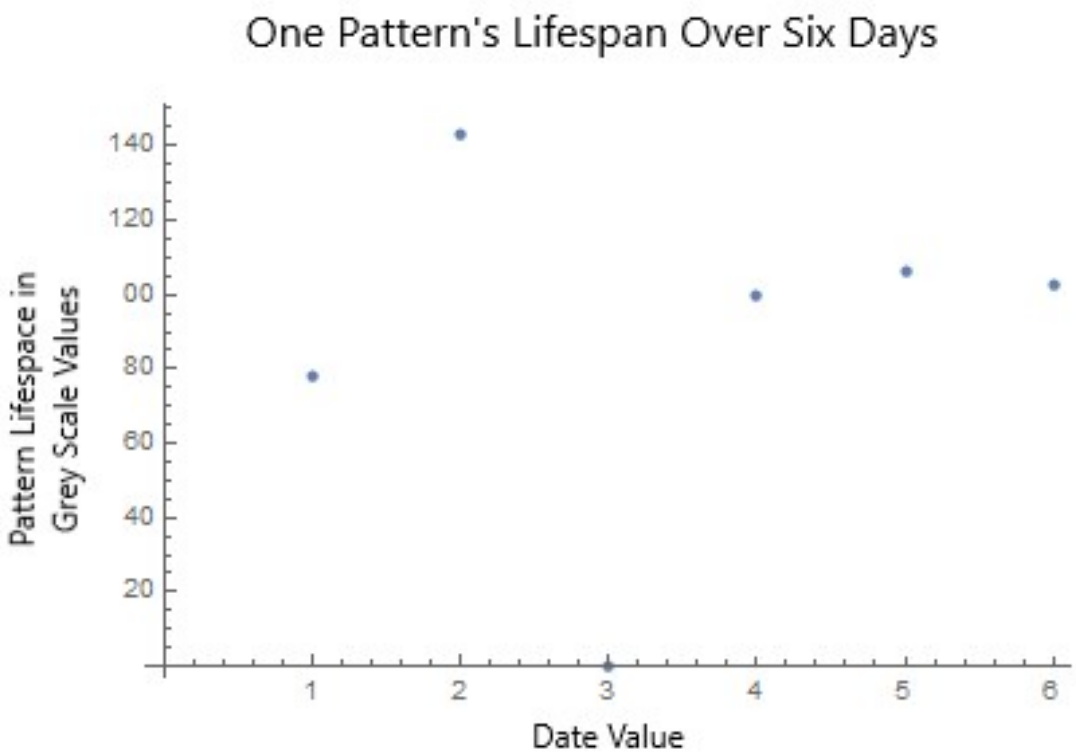



## Convertio Converted & Blacked Out



# Comparison of Lifespan of Persistent Pairs

Comparing lifespan of first persistent pairs for the 6 dates we get following diagram:



# Future Directions

1. Different heat maps require different RGB-to-grayscale conversions. We discussed 3 methods, but further techniques and customizability may be necessary.
2. More computational resources would allow us to execute predictive analysis of crime data.
3. The end goal would be to develop a tool that takes as input any collection of heat maps (from distribution of resources, housing prices, crime statistic, ... ) and is able to use barcodes to predict future patterns.

# Conclusion & Links

Questions?

Links:

1. **Paper**: https://faculty.math.illinois.edu/∼rasekh2/dmt_vf_ms.pdf
2. **Slides**: https://faculty.math.illinois.edu/∼rasekh2/jmm.pdf
3. **Code**: https://github.com/redavids/IBTCDA/tree/master